REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-01-

0160

1. AGENCY USE ONLY (*Leave Blank*)

2. REPORT DATE
20 Feb 01

3. REPORT TYPE AND D...
Final 1 Dec 98 - 30 Nov 00

4. TITLE AND SUBTITLE
Near-Field and Distance Cues in Auditory Spatial Displays

5. FUNDING NUMBERS
G F49620-98-1-0108

6. AUTHORS
Barbara G. Shinn-Cunningham
Douglas Brungart

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Boston University
Department of Cognitive and Neural Systems
677 Beacon St., Boston, MA 02215

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
801 N. Randolph St.
Arlington, VA 22203

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Analysis showed that, when sources are within a meter of the listener, binaural differences resolve source location to within a torus of space. Localization studies demonstrated that despite the fact that the direct sound reaching the ears is relatively intense for nearby sources, the reverberation dramatically improves distance perception. In fact, it also causes minor degradations in directional localization accuracy; however, these effects are relatively minor. Results also suggest that in a room listener's accuracyat judging source distance and direction improved gradually with time, even in the absence of direct feedback about localization performance.

Headphone studies confirmed that reverberation is a dominant cue for source distance, even for nearby sources in which the reverberant energy is low and binaural cues provide distance information. Results suggest that distance is computed primarily from the signal reaching the nearer ear and is a primarily monaural cue. Thus, in order to simulate distance in a spatial auditory display, including realistic reverberation is critical.

Acoustical measurements demonstrate that many features of the signals reaching a listener's ears change with distance and with reverberation. Further work must be performed in order to determine which of these features is critical in the perception of source distance.

14. SUBJECT TERMS
Spatial Hearing
Acoustic Displays
Distance Perception

15. NUMBER OF PAGES
11

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
UNCLASSIFIED

18. SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

19. SECURITY CLASSIFICATION OF ABSTRACT
UNCLASSIFIED

20. LIMITATION OF ABSTRACT
UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-1
298-102

Final Report

Near-Field and Distance Cues in Auditory Spatial Displays

EXECUTIVE SUMMARY

This report summarizes the results of the collaboration between researchers at Boston University and the Air Force Research Laboratory Human Effectiveness Branch investigating the acoustics and psychoacoustics of sound localization for sound sources near a listener's head. The results of this work are unique in that little previous work has examined how the acoustics of the signals reaching a listener change with source distance. This work is critical for determining how to robustly represent sound source distance in spatial auditory displays.

Theoretical analysis showed that the interaural differences that arise in anechoic space when sources are within a meter of the listener resolve source location to within a torus of space (in contrast with previous studies of more distant sources, where it has long been known that binaural cues are roughly constant for a source on a cone centered on the interaural axis). Localization studies in a reverberant room demonstrated that despite the fact that the direct sound reaching the ears is relatively intense compared to the reverberation for nearby sources, the reverberation dramatically improves distance perception. In fact, it also causes minor degradations in directional localization accuracy; however, these effects are relatively minor. Surprisingly, results also suggest that in a room, unlike in anechoic space, a listener's accuracy in judging source distance and direction improved gradually with time, even in the absence of direct feedback about localization performance.

Headphone simulation studies confirmed that reverberation is a dominant cue for source distance, even for nearby sources in which the reverberant energy is not very intense and binaural cues can provide some distance information. In fact, our results suggest that distance is computed primarily from the signal reaching the nearer ear and is a primarily monaural (not binaural) cue. Thus, in order to simulate distance in a spatial auditory display, including realistic reverberation is critical; however, we cannot yet say exactly what aspect of reverberation provides this information.

Acoustical measurements demonstrate that many features of the signals reaching a listener's ears change with distance and with reverberation; further work must be performed in order to determine which of these features is critical in the perception of source distance.

Finally, while including reverberation is undoubtedly helpful for generating a realistic percept of acoustic sources in space, it may cause degradations on other important tasks. Preliminary studies investigating how spatial separation of competing sources affects detection and speech intelligibility in anechoic simulations show that large changes in threshold signal levels are observed when nearby target and masker sources are moved in space. The way in which these thresholds are affected by reverberation must be evaluated in order to determine how best to trade off spatial accuracy of a display against performance on other tasks.

PERSONNEL

The personnel who were supported by and/or associated with this project are Principal Investigator Barbara Shinn-Cunningham (Boston University), Co-Principal Investigator Douglas Brungart (Armstrong Laboratory), consultant Nat Durlach (Massachusetts Institute of Technology), graduate students Fevzi Alimoglu (B.U.), Tara Brown (M.I.T.), Norbert Kopco (B.U.), Sergey Lubensky (M.I.T.), and Scott Santarelli (B.U), undergraduates Paul Ilardi (B.U.), Arpan Jhaveri (B.U.), Lisa Mraz (B.U.), Jason Rodriguez (B.U.), Jason Schickler (B.U.), and James St. Pierre (B.U), and collaborators Abhi Kulkarni (Bose Acoustics) and Bill Rabinowitz (Bose Acoustics).

PUBLICATIONS

Shinn-Cunningham, B. G. (1998). "Applications of virtual auditory displays," 20th Ann Conf IEEE Eng Med Biol Soc, Hong Kong, China, 29 October – 1 November, **20**, 1105-1108.

Brungart, D. S. (1998). "Preliminary model of auditory distance perception for nearby sources" NATO ASI on Auditory Computational Hearing, Il Ciocco, Italy.

Brungart, D. S. (1998). "Control of perceived distance in virtual auditory displays," 20th Ann Conf IEEE Eng Med Biol Soc, Hong Kong, China, 29 October – 1 November, **20**, 1101-1104.

Brungart, D. S. (1998). "Three-dimensional auditory localization of nearby sources" 42nd Annual Meeting of the Human Factors and Ergonomics Society.

Brungart, D. S. and W. M. Rabinowitz (1999). "Auditory localization of nearby sources I: Head-related transfer functions." Journal of the Acoustical Society of America **106**(3): 1465-1479.

Brungart, D. S. and N. I. Durlach (1999). "Auditory localization of nearby sources II: Localization of a broadband source in the near field." Journal of the Acoustical Society of America **106**(4): 1956-1968.

Brungart, D. S. (1999). "Auditory localization of nearby sources III: Stimulus effects." Journal of the Acoustical Society of America **106**(6): 3589-3602.

Kramer, G. Walker, B., Bonebright, T., Cook, P., Flowers, J., Miner, N., Neuhoff, J., Bargar, R., Barrass, S., Berger, J., Evreinov, G., Fitch, W. T. Gröhn, M., Handel, S., Kaper, H., Levkowitz, H. Suresh L., Shinn-Cunningham, B. Simoni, M., Tipei, S. (1999).*Sonification Report: Status of the Field and Research Agenda*. Palo Alto, CA.

Shinn-Cunningham, B. G. (2000). "Adapting to remapped auditory localization cues: A decision-theory model." Perception and Psychophysics, **62**(1), 33-47.

Shinn-Cunningham, B.G., Santarelli, S., and Kopco, N. (2000). "Tori of confusion: Binaural cues for sources within reach of a listener," J Acoust Soc Am, **107**(3), 1627-1636.

Shinn-Cunningham, B.G. (2000). "Learning reverberation: Considerations for spatial auditory displays," in Proceedings of the International Conference on Auditory Display, Atlanta, GA, 2-5 April 2000, 126-134.

Shinn-Cunningham, B. G. (2000). "Distance cues for virtual auditory space," Proceedings of the IEEE 2000 International Symposium on Multimedia Information Processing, Sydney, Australia, 13-15 December, 2000.

Brungart, D. S., W. Rabinowitz and N. I. Durlach (2000). "Evaluation of response methods for near-field auditory localization experiments." Perception and Psychophysics **62**(1): 47-63.

Shinn-Cunningham, B. G. (2001). "Spatial auditory displays," in *International Encyclopedia of Ergonomics and Human Factors*, W. Karwowski (ed.). London: Taylor and Francis, Ltd. (in press).

Shilling, R. D., & Shinn-Cunningham, B. G. (2001). "Auditory channel," in Handbook of Virtual Environments Technology, K. Stanney (ed.). (in press).

Shinn-Cunningham, B. G., J. Schickler, N. Kopco and R. Y. Litovsky (2000). "Spatial unmasking of nearby speech sources in a simulated anechoic environment." J Acoust Soc Am (under revision).

Brungart, D.S. Near-Field Virtual Audio Displays (2000). Presence (submitted).

RESULTS

1. Localization Cues

In order to improve our theoretical understanding of the acoustic signals that arise at the ears when sources are within reach, we have analyzed and modeled how spatial cues change with position for sources within a meter of the head. We have also measured spatial acuity in this region in a series of headphone studies measuring the just-noticeable difference (JND) in azimuth as a function of distance for nearby sources. These studies show that spatial acuity is substantially improved for very close sound sources (<19 cm) directly in front of the head, but that there are no consistent variations in the JND with distance for other directions. However, results indicate that monaural azimuthal JNDs (measured at the ipsilateral ear) are substantially improved for nearby sources, probably due to the large overall level variations with azimuth that occur when the source is near the head.

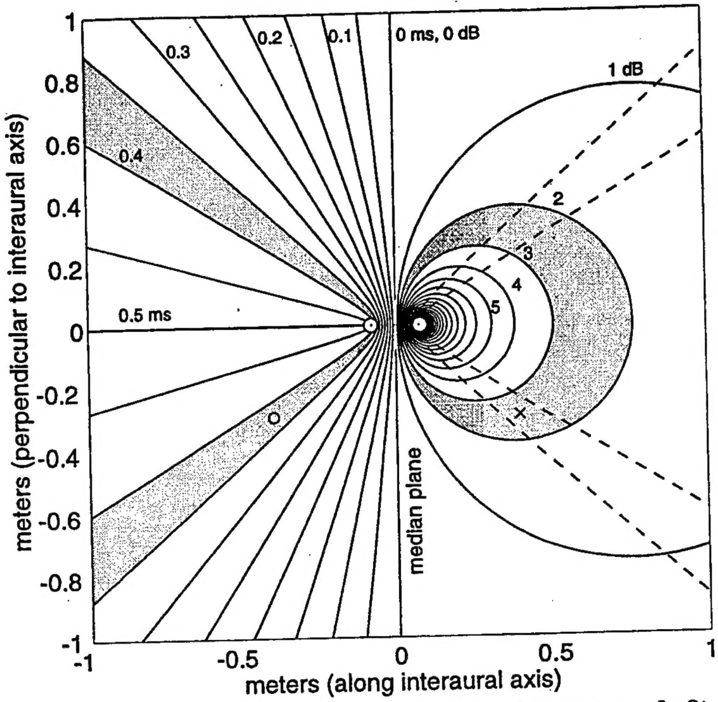

Figure 1: Iso-ITD (every 50 microsec; left side) and ILD (every 1 dB; right side) showing source positions in a plane containing the ears that would give rise to identical interaural differences at the ears (denoted by small open circles). Gray area shows region of space that a source at the 'o' (left side) or 'x' (right side) could occupy while giving rise to approximately the same interaural difference.

In the simplest theoretical analysis of binaural localization cues, the acoustic effects of the head itself were ignored (an "acoustically-transparent head" condition) and the loci of points for which all sources give rise to the same interaural differences of time or level were computed algebraically. While the resulting iso-ITD contours have been discussed in the literature for decades (e.g., see the description of cones of confusion in von Hornbostel and Wertheimer, 1920), surprisingly, this analysis had never before been undertaken (to our knowledge) for interaural intensity differences. In three space, iso-IID surfaces are perfect spheres whose centers fall along the interaural axis. As a result, the locus of points for which sources give rise to the same ITDs and IIDs is a circle centered on the interaural axis (see Figure 1). If one further assumes that there is some uncertainty in binaural perception, this circle is "smeared" into a finite volume: a torus (a solid of rotation) centered on the interaural axis. The volumes of these "tori of confusion" change with source position; the torus degenerates to the normal cone of confusion for distant sources (where IID changes slowly with position) and to the median plane for sources far from the interaural axis (where IID cues are uniformly zero).

Of course, this analysis ignores the effects of the head. A more complete mathematical analysis (based on iteratively calculated numerical solutions) was performed in which the head is treated as a rigid sphere. This analysis was repeated for ears located at diametrically opposite points on the surface of the sphere as well as when the ears are displaced backward (more like the ears on a listener's head). The resulting ITD values changed very little in perceptual terms when the effects of the head were included, increasing only slightly for sources very near one ear (e.g., see Shinn-Cunningham, Santarelli, and Kopco, 2000). However, IIDs increased dramatically in this analysis compared to solutions for a "acoustically transparent head." In addition, IIDs for a rigid spherical head depend on frequency; at low frequencies the

"acoustically transparent head" model is relatively accurate, but the total IID increases dramatically for high audible frequencies. Analysis also showed a surprising result: the total IID for a source at a particular position can be factored into two essentially additive components. The first is a frequency- and direction-dependent component that is independent of distance (the normal "head shadow"). The second depends only on the relative distance from the source to the two ears (i.e., is the component derived for the "acoustically transparent" head) and is roughly independent of frequency. Thus, taken together, IID and ITD information should be sufficient to determine source location to within a torus of confusion (equivalent to that described by the "acoustically transparent head" analysis). (See Shinn-Cunningham et al., 2000).

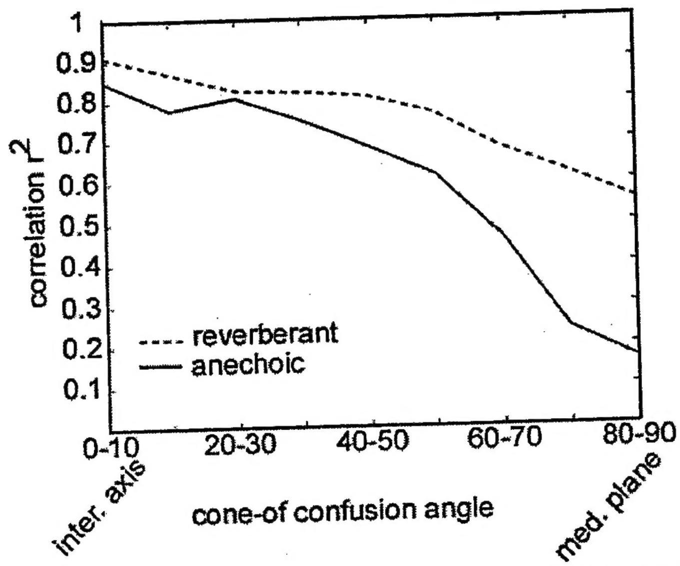

Figure 2: Correlation (r^2) between source and response distance as a function of source cone of confusion angle in anechoic conditions (solid) and reverberant conditions (dashed lines), averaged across subject.

New analysis of free-field localization data for nearby sources (Brungart and Durlach, 1999) supports the torus of confusion analysis. In this earlier study (which inspired much of the work performed in the current grant), subjects were asked to localize sources presented at a randomly roved overall level from locations in the right hemifield by pointing to the heard position of the source. When these data are analyzed in terms of θ, the cone-of-confusion angle (corresponding to the ITD); ϕ, the angle around the interaural axis (corresponding to the spectral cue); and r, distance (which is partially correlated with IID), patterns of localization accuracy are easy to explain. Most response bias and response variability can be ascribed to errors in ϕ (which is calculated from the relatively ambiguous spectral cue) and r (for which robust cues are not uniformly available in all regions of space). In contrast, errors in θ are uniformly small. If one analyzes the errors in units of ITD and IID (by converting source and response locations to corresponding binaural cue values using the "acoustically-transparent head" model), the magnitude of average binaural errors are roughly the size of JNDs found in discrimination tasks. The average absolute ITD error is only 53 μs; the average absolute IID error is 1.1 dB (see Shinn-Cunningham, 2000). Analysis of distance perception as a function of angle from the interaural axis is also instructive (see Figure 2). The main distance cue for nearby sources in anechoic space is the IID. Although the IID varies dramatically with distance for sources near the interaural axis (providing a rich distance cue for sources in this part of space), it is approximately zero, independent of distance for sources near the median plane. The corresponding behavioral result can be seen in Figure 2, which plots the square of the correlation coefficient between source and response distances (on a logarithmic scale) as a function of θ for the anechoic study (solid lines). In Figure 2, one sees a progressive decrease in distance performance as the source approaches the median plane and the IID cue becomes less salient.

We have also made some progress in understanding how the spectral localization cues related to the directional properties of the pinna vary with the distance of a nearby sound source. In this analysis, KEMAR-manikin HRTFs for nearby sources were used to map the relationship between azimuth and distance for high-frequency HRTF features (> 8 KHz) associated with the directional filtering properties of the outer ear. Specifically, we determined the azimuths of the HRTFs measured at distances of 12 cm, 25 cm, and 50 cm that most closely matched the high-frequency features of far-field HRTFs measured at 0, 30, 60, 150, and 165 degrees in azimuth. The results show that the high-frequency HRTF features vary

with the location of the source relative to the ear rather than relative to the center of the head, resulting in an auditory parallax effect for nearby sources (Brungart, 1999). This finding is relevant to the construction of three-dimensional auditory displays because it suggests that individualized HRTFs may be produced for source locations near the head by geometrically warping a single set of far-field HRTFs. This would mitigate the technical problems associated with HRTF measurements near the head.

2. Localization in Reverberant Space

In order to begin to quantify how reverberation affects localization performance in reverberant space, an earlier free-field localization study (Brungart and Durlach, 1999) was replicated in a reverberant room (using a randomly-roved presentation level). The most important finding was that even when sources are near a listener and the relative level of reverberation is low, distance performance is significantly better than in anechoic space. This can be seen in Figure 2 (dashed line), which compares distance perception in reverberant and anechoic conditions (plotted as the square of the correlation coefficient between source and response distance as a function of the cone of confusion angle of the source). Whereas distance perception degrades as the source approaches the median plane in anechoic space, performance in reverberant space is better and varies less dramatically with source direction. This result demonstrates that reverberation provides a robust cue for source distance. Of course, these results by themselves do not address whether the reverberation cue provides distance information in addition to, or instead of, the distance information in the IID. Specifically, since even in the reverberant condition, distance performance degrades as the source approaches the median plane, these results leave open a number of intriguing questions. What aspect of reverberation provides the cue for distance? Is the IID cue the dominant cue for sources along the interaural axis in reverberant space? Does the reverberant cue itself degrade as sources approach the median plane?

Another smaller, but no less interesting, difference between anechoic and reverberant localization performance was also evident. Localization accuracy of subjects in reverberant space gradually improves over time. This long-term learning (which is not seen in anechoic results) occurs in every spatial dimension examined, including in directional localization. This effect is described in more detail in Shinn-Cunningham, 2000b). Apparently, the small distortions of localization cues caused by reverberation can cause small errors in localization, but these errors gradually diminish with experience in a particular room.

In sum, even when reverberation is relatively quiet (i.e., for sources near the listener), reverberation provides robust information about source distance and causes measurable degradations in directional localization. With practice in the room, localization accuracy shows statistically significant improvements; however, the absolute magnitude of these effects is small.

In another study, we tested whether localization in distance and direction would be worse in a reverberant room when the envelope of the direct sound rose and decayed slowly over time (100 ms-long cosine-squared envelope). We hypothesized that localization might benefit from sharp onsets, since, at these transition times, the listener would get a good "look" at the direct sound, undistorted by reverberation (or, in the case of the offset, would get a good look at the reverberation decay pattern in isolation). In turn, such looks might allow the listener to estimate the relative energy in the direct sound versus the reverberation more accurately and allow accurate distance perception. However, including slow envelope onsets (and offsets) in the direct sound did not cause any significant changes in localization performance in a room. In other words, it appears that the reverberant distance cue does not depend on the presence of rapid temporal modulations of the source signal.

3. Robust Simulation of Distance Cues

In order to begin to tease apart how reverberation encodes source distance, a headphone-based experiment was performed in which both anechoic and reverberant distance perception could be compared directly. In order to perform this study, we wanted to be able to create realistic simulations of sources at different directions in both anechoic and reverberant conditions using headphones.

Measurements of the HRTFs in the reverberant room used in the real-world localization experiments were taken at seven different distances (15, 19, 25, 38, 50, 75, and 100 cm) at lateral positions to the right side of the head (along the interaural axis) and directly in front of the listener. A Golay code technique was used to present pink noise stimuli to the ears of each experimental subject. The resulting HRTFs (derived from the signals measured at the ears) contained both a direct HRTF (essentially the typical "anechoic" HRTF for a source at the position of the speaker) and reverberant energy (the superposition of many sources arriving from many different directions, coming from the floor, ceiling, and walls). For all these recordings, the anechoic portion of the HRTF had dissipated before the first reflection (from the floor) arrived. As a result, these HRTFs could either be time-windowed to generate pseudo-anechoic simulations or used in full to generate reverberant-room simulations.

The measured "anechoic" and reverberant HRTFs were convolved with noise samples (identical to those used in the real-room localization experiments) to generate headphone simulations of sources along the interaural axis or in the median plane at various distances. As in the real-world localization task, overall presentation level was randomly roved to remove intensity as a distance cue. Stimuli were presented both monaurally (where the signal to the left earphone was turned off) or binaurally. Trials were blocked by condition (anechoic monaural, anechoic binaural, reverberant monaural, and reverberant binaural) with blocks presented in random order. No feedback was provided to the subjects during the experiment.

We hypothesized that in the anechoic monaural condition, subjects would be poor at judging distance, but that in the anechoic binaural condition, subjects would be good at distance judgements for lateral sources and poor for medial sources. If reverberant distance cues were primarily monaural, then the reverberant monaural and reverberant binaural results would be similar for the medial sources. However, since reverberation will decorrelate the signals at the two ears, performance might be generally better for reverberant binaural than for reverberant monaural presentations. Finally, if reverberant binaural results were better than reverberant monaural sources for lateral but not medial sources, it would indicate that the reverberation cue was primarily monaural, but was used in addition to the binaural IID cue.

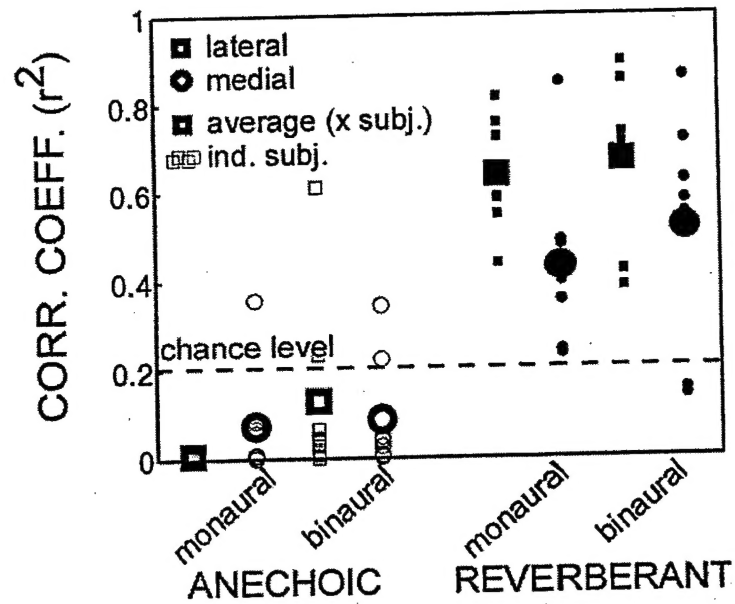

Figure 3: Correlation (r^2) between source and response distance in headphone experiments. Anechoic results (left) are generally below chance; reverberant results (right) are generally good. For reverberant results, lateral results (squares) are better than medial (circles) and monaural results are comparable to binaural.

Figure 3 plots the square of the correlation coefficient between source and response distance (separately for each condition and for lateral and medial sources). In fact, the most obvious finding was that the anechoic results (left side of the figure) resulted in uniformly poor distance judgements, with the distance correlation failing to be significantly stronger than chance (within a 95% confidence interval). This result held even for binaural presentations of lateral sources, where there was a 15 dB change in the IID with distance for the simulated sources.

Looking at the reverberant results, one can see that performance for lateral sources is better than for medial sources for both monaural and binaural conditions. This implies that the reverberant cue for distance of nearby sources is stronger for lateral than medial sources. Examining only the lateral sources (where the binaural IID cue is strong), there is essentially no difference in distance performance between monaural and binaural conditions; in other words, the binaural distance cue does not appear to improve distance judgements. Finally, there is a small improvement in distance judgements for medial reverberant sources when they are presented binaurally rather than monaurally. This final result is somewhat surprising, since there is essentially no binaural distance information for medial sources. However, examination of the mean responses for source distance in these conditions provided further insight into what occurred.

For the lateral sources in reverberant conditions, distance judgements were not only highly correlated with the source distance for both binaural and monaural presentations, the mean judgements showed essentially no bias. However, for the medial sources in the reverberant conditions, mean judgements of source distance were consistently farther from the head then the actual source distance for binaural stimuli. This bias increased significantly for monaural stimuli. We believe that the perceived direction of the source influenced how an essentially monaural reverberation cue for distance is interpreted. When "medial" sources are presented monaurally, the perceived direction is along the interaural axis (in other words, binaurally- and monaurally-presented medial sources are perceived in different directions). In contrast, binaurally-presented lateral sources are already heard in the direction of the interaural axis, and turning off the far ear has no effect on the perceived direction. Finally, the direct-to-reverberant energy ratio (which has been suggested as a cue for distance) varies with direction as well as with distance for nearby sources. In fact, our analysis shows that a source at a particular distance at a lateral location gives rise to a larger direct-to-reverberant energy ratio than a source at the same distance directly in front of the listener. While previous studies have not shown an interaction between perceived direction and perceived source distance in reverberant space, this direction dependence is only pronounced for sources very close to the head, like those in the current study.

To summarize these findings, we conclude that reverberation is a much more salient and robust distance cue than the anechoic IID cues in headphone simulations. While the IID undoubtedly has an effect on perceived direction in reverberant environments (e.g., see Hartmann, Constan, and Rakerd, 1999; Hartmann and Rakerd, 1999), the unique, distance-dependent IIDs appear to be relatively unimportant for distance perception when listening under headphones. The reverberant distance cue is essentially a monaural cue; however, perceived source direction influences how the reverberant cue (whatever form it takes) is mapped to perceived source distance.

A follow-up experiment was performed to ensure that the large IIDs in the anechoic binaural simulations of lateral sources were perceptible. In this experiment, subjects performed a seven-alternative forced-choice task in which only the seven lateral sources were presented binaurally at random levels, but correct-answer feedback was provided. Subjects were repeatedly tested until performance stabilized. The results of this experiment confirmed that the large modulation of IID with source distance is perceptible (performance for each subject improved substantially, with the final performance of five of the six subjects substantially better than chance and the sixth subject's performance approaching chance). We believe that anechoic cues can convey distance information for lateral sources in anechoic space; however, these cues are not as robust or as compelling as reverberant distance cues, at least for headphone simulations.

4. Models of Distance Perception

A simplified model of auditory distance perception in anechoic space was developed for sound sources near the head. This model assumes that listeners can perfectly determine the left/right position of a sound source from the ITD, but sometimes confuse sources in the front and rear hemispheres. They then use the estimated IID (which is assumed to be a Guassian-distributed random variable) to estimate distance for the perceived azimuth location. The perceived distance equals the actual distance that would produce the IID

matching the listener's perceptual estimate. The model can account both the mean responses and variation in the responses that were observed behaviorally. Simulations show that the model can predict performance reasonably accurately for lateral sources more than 30 degrees from the median plane; however, the model underestimates performance in the median plane (where no non-binaural cues are available to listeners). This model makes a number of simplifying assumptions. For instance, the model assumes that ITD and IID are separately available to the listener and that ITD alone determines perceived direction (i.e., there is no time-intensity trading of any sort). The model also assumes that source direction is perceived perfectly (except for possible front/back reversals). Of course, we do not believe that these assumptions are entirely correct. We were simply trying to explore whether we could model the accuracy of distance perception if we assumed that the listener was able to extract reliable directional information (either from ITD or from a combination of ITD and IID cues) and then used the IID to determine distance.

In reverberant space (and reverberant simulations), it appears that distance perception improves due to some monaural attribute of the total signal reaching the ears. Only one computational model for reverberant distance perception has been proposed (Bronkhorst and Houtgast, 1999). However, this model assumes that sources are relatively far away so that the intensity of the direct sound varies inversely with the square of the source distance. For sources relatively close to the head, the direct sound intensity increases more rapidly than this. In fact, the direct sound intensity increases more rapidly for lateral than for medial sources (i.e., there is a directional dependence of the direct-to-reverberant energy ratio for nearby sources). If one were to refine the proposed model of distance perception so that distance judgements are based on the appropriate direct-to-reverberant energy ratio for sources in a particular (perceived) direction, it would predict the direction of the response bias seen for monaurally-presented medial sources reported in the previous section. However, our analysis shows that even a refined version of the model fails to predict the magnitude of the bias we observed. In particular, the bias should be even larger than was found, even when the model is liberally "doctored" to fit the acoustics of nearby sources and to take into account direction-dependent changes in the levels of the direct sound with distance. Finally, as pointed out in the background section, this model is not plausible, since it assumes that listeners can perfectly deconvolve room and HRTF characteristics (a feat that is impossible even in anechoic space; see Rakerd, Hartmann, and McCaskey, 1999).

As a result of our analysis, we have looked for alternative statistics that could explain distance perception for nearby sources in reverberant space. As yet, we have not found any alternatives that can account for all aspects of our results, but we have ruled out binaural correlation, monaural autocorrelation, and variants of the Bronkhorst and Houtgast model that operate on the energy in the total signal reaching the ears rather than on the energy in the reconstructed room HRTF.

5. Predictions of Spatial Unmasking in Anechoic Space

We have made substantial progress is in analyzing what "natural" combinations of binaural cues can arise for sources within reach of a listener. In particular, from the analysis of "tori of confusion," we can predict both interaural phase and intensity differences for sources close to the head. This analysis is particularly relevant for understanding how spatial unmasking effects may play out for nearby target and masker sources. Our analysis points out a number of interesting aspects of the conditions that can arise for nearby sources. Of course, one of the major effects of presenting sources very close to the listener is that the overall energy received at the ears changes dramatically with source distance.

We have analyzed how the "received energy" (the monaural contribution to spatial unmasking) changes with source location for nearby sources. This analysis, based on the predictions of the spherical model of the head, shows that this monaural component contributes significantly (and for some spatial configurations overwhelmingly) to the observed changes in the detection threshold.

We have also looked at the more interesting binaural effects on spatial unmasking. We have performed simple analyses (based on the Colburn model of binaural processing; e.g., see Colburn, 1973) to predict

changes in detection threshold after normalizing the SNR at the so-called "better ear." Previous studies have generally assumed that the better ear is the ear closer to the source (an assumption that is generally valid for relatively distant sources). However, as our analysis for nearby sources shows, a prediction of which ear is the better ear must take into account the relative distances of the target and the masker to both ears. For example, if a masker is very close to the right ear (along the interaural axis) and the target is farther away (but in the same direction), the left ear will be the better ear even though the target source is directly to the right side of the head. This better-ear advantage is very small when the masker is more than a meter away (i.e., there is no better ear); however, the advantage is pronounced for nearby sources. Predictions for which ear is the better ear are easily understood from analysis of IID cues. For a given masker location, any target source on the same iso-IID surface will yield the same SNR at both ears (i.e., once again, there is no better ear). If the target source gives rise to a different IID than the masker, the better ear (that with the better SNR) is the ear for which the IID is bigger than for the masker. In other words, if the masker yields an IID of 10 dB favoring the right ear and the target yields an IID of 8 dB favoring the right ear, the left ear has a better signal to noise ratio than the right (by 2 dB). Of course, the actual IIDs that arise for nearby sources vary dramatically with frequency. Therefore, it is very likely that the better ear at one frequency is not the better ear for all frequencies. Combining these observations with the standard Colburn model (Colburn, 1973) allows one to predict detection thresholds for narrowband targets at various locations relative to a wideband masker. For many target/masker geometries, the predicted amount of binaural gain (i.e., the unmasking observed after equating the SNR at the better ear) is small compared to the monaural effects. But for all spatial configurations studied with a 500Hz target, the binaural gain is always significant (3-10dB).

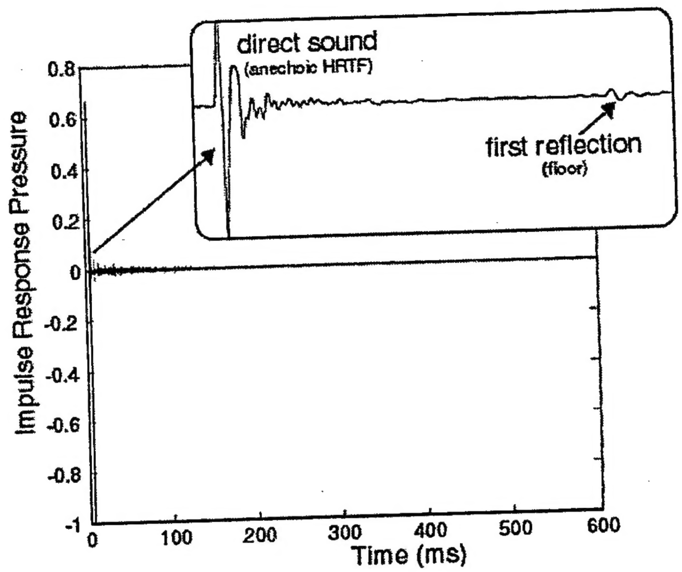

Figure 4: Time-domain impulse response at the ear of a listener for a source in a room positioned 1 meter from the head at ear level, in the right front quadrant. Insert shows initial 25 ms of response, including direct sound and first reflection

6. Preliminary Measurements of HRTFs in Reverberant Space

We have begun to analyze the acoustic effects of reverberant energy on the total signals reaching the ears in order to identify and test alternative hypotheses for how distance is computed from reverberant cues. Based on feedback from other researchers in the field, we decided to switch from using Golay codes to using Maximum Length Sequences (e.g., see Rife and Vanderkooy, 1989; Vanderkooy, 1994) in order to get more robust, reliable measures of room HRTFs.

Results of analysis of our individualized HRTF measurements using the MLS approach are quite encouraging. Figure 4 presents a sample HRTF in which both the total (large axes) and the anechoic portion (insert) of the HRTF for one ear are shown for an example subject and location.

Many acoustic features vary with distance and direction of a sound source near the listener, and might be able to predict the behavioral results observed. For instance, in Figure 6, we show the direct sound energy relative to the reverberant energy, like the statistic proposed by Bronkhorst and Houtgast (1999) to underlie distance judgments. Results are showna as a function of source distance, for two source

directions at both the left and right ears. The direct-to-reverberant energy ratio does provide a distance cue. However, it should be noted that nearly any other attribute of the signals reaching the ears that vary with reverberation with be correlated with the changes in the direct-to-reverberant energy ratio.

We performed analysis comparing the pseudo-anechoic portion of the HRTF to the total HRTF in reverberant space. The reverberation causes near-random fluctuations in the total spectrum of the signals reaching the ears around their "true" anechoic values. The size of these fluctuations varies with source distance, as one might expect. We are encouraged by our observations for a number of reasons. This property of the reverberant spectrum will be highly correlated with the direct-to-reverberant energy ratio and should be able to account for distance perception of sources beyond a meter from the listener (i.e., those positions that are correctly predicted by Bronkhorst and Houtgast, 1999). In addition, computation of spectral "roughness" is computationally feasible, unlike the proposed model (Bronkhorst and Houtgast, 1999). In the upcoming year, we intend to explore various estimates of spectral roughness as a cue for source distance in reverberant rooms.

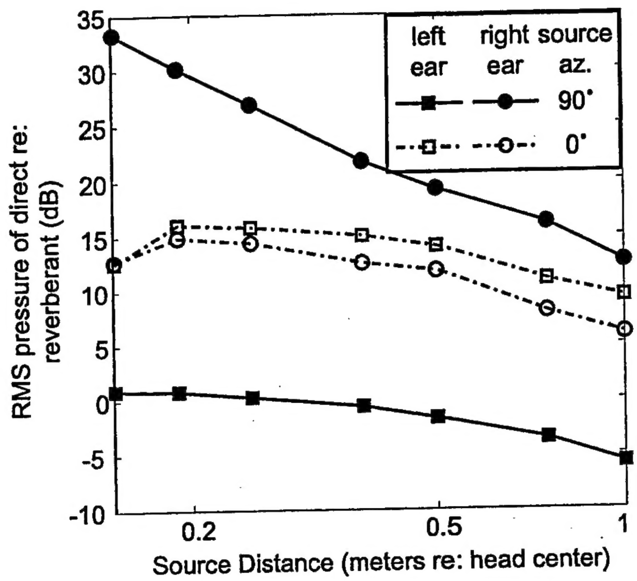

Figure 6: Ratio of direct to reverberant RMS pressure level at left (square) and right (circle) ears as a function of distance for sources straight ahead (dashed lines) and 90° to the right (solid lines).

7. Distance Perception of Speech Stimuli

In addition to the studies of localization of unfamiliar sources in anechoic and reverberant space already reported, we performed some preliminary studies examining distance perception of speech signals. Speech is a unique stimulus for distance perception because the acoustic properties of the speech change consistently with the output level of the talker. Consequently, it is possible to estimate the loudness of a live talker based on the apparent vocal effort of the speech signal. By comparing this apparent level of vocal effort to the level of the stimulus reaching the ears, it is possible to estimate the distance of the talker. A listener hearing a loud whisper, for example, knows the sound must originate from a nearby talker, while a quiet shout must originate from a distant talker. The results of these experiments demonstrate that vocal effort is an extremely potent cue for absolute distance. However, these experiments are substantially different in character from the others topics discussed in this proposal and will not be pursued further in this research program.

REFERENCES

Bronkhorst, A. W. and T. Houtgast (1999). "Auditory distance perception in rooms." Nature **397**(11 February): 517-520.

Brungart, D. S. (1999). Auditory parallax effects in the HRTFs for nearby sources. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York.

Brungart, D. S. and N. I. Durlach (1999). "Auditory localization of nearby sources II: Localization of a broadband source in the near field." Journal of the Acoustical Society of America **106**(4): 1956-1968.

Colburn, H. S. (1973). "Theory of binaural interaction based on auditory-nerve data. I: General strategy and preliminary results on interaural discrimination." Journal of the Acoustical Society of America **54**: 1458-1470.

Hartmann, W. M., Z. A. Constan, and B. Rakerd (1999). "Binaural coherence and the localization of sound in rooms." Journal of the Acoustical Society of America **103**(5): 3081.

Hartmann, W. M. and B. Rakerd (1999). "Localization of sound in reverberant spaces." Journal of the Acoustical Society of America **105**(2): 1149.

Rakerd, B., W. M. Hartmann, and T. L. McCaskey (1999). "Identification and localization of sound sources in the median sagittal plane." Journal of the Acoustical Society of America **106**(5): 2812-2820.

Rife, D. D. and J. Vanderkooy (1989). "Transfer-function measurement with maximum-length sequences." Journal of the Audio Engineering Society **6**: 419-444.

Shinn-Cunningham, B. G. (2000). Learning reverberation: Implications for spatial auditory displays. International Conference on Auditory Displays, Atlanta, GA.

Shinn-Cunningham, B. G., S. Santarelli, and N. Kopco (2000). "Tori of confusion: Binaural localization cues for sources within reach of a listener." Journal of the Acoustical Society of America **107**(3): 1627-1636.

Vanderkooy, J. (1994). "Aspects of MLS measuring systems." Journal of the Audio Engineering Society **42**(4): 219-231.

von Hornbostel, E. M. and M. Wertheimer (1920). Uber die Wahrnehmung der Schallrichtung [On the perception of the direction of sound]. Sitzungsber. Akad. Wiss. Berlin: 388-396.